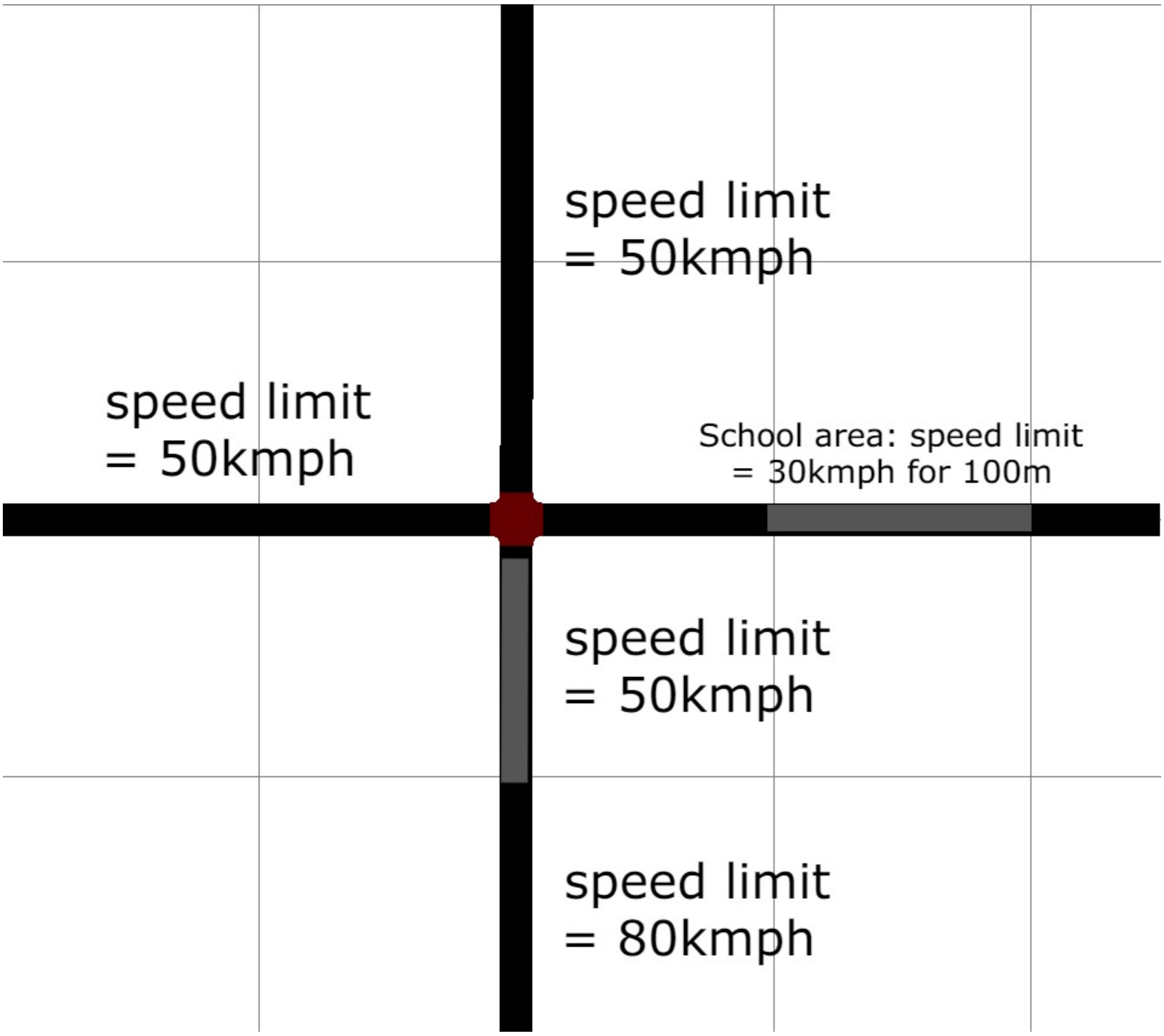
Weekly Updates (30/11/23) – Vinayak Gajendra Panchal

Ego vehicle collision prediction

Simulation setup:

Step-length	0.10 sec (10 cycles/sec)
Collision.action	remove
Collision.stoptime	2 sec
Collision.check-junctions	True
'- - log' (simulation log files)	simulation_main.log
Number of Vehicles	11 (1 ego-vehicle, 10 non-ego vehicle)

2-Lane 4-Way Junction Details:



Dataset Creation: I have integrated lanes with speed limits, including a school zone with a maximum speed of 30 km/h (marked in grey), to better replicate traffic scenarios in the LaneLet2 framework. In this simulation, 11 vehicles were modeled, with data collected every 0.1 seconds across three different sets (training, validation, and testing). This produced three comprehensive datasets, each with Full Cycle Data (FCD) points and 27 unique attributes. However, for predicting collisions, these datasets may not be directly applicable, as they encompass data for all vehicles at various timestamps and are not arranged as time-series data, whereas our focus is on the ego-vehicle and its specific timestamps. Furthermore, not all the attributes are necessary for collision prediction. Therefore, our initial step is to extract a subset of the dataset, focusing only on the timeframe when the ego-vehicle is active and capturing data for all vehicles present during that period.

To make this dataset suitable for the prediction of collision, I aggregated the data to capture key features of ego-vehicle (for all train-validation-test dataset):

1. **Time:** Time in seconds when ego vehicle was present in the simulation.
2. **ego-speed, ego-acceleration: The** acceleration and speed value of ego-vehicle at all timestamps
3. **ego_x, ego_y:** spatial data of ego-vehicle at all timestamps.
4. **non_ego_flow.X_relative_speed:** Calculated relative speeds of all non-ego vehicles at each timestamp with the ego-vehicle. Here 'X' in non_ego_flow.X represents non-ego vehicle number.
5. **non_ego_flow.X_relative_displacement:** Calculated relative displacement (Euclidian distance) of all non-ego vehicles at each timestamp with the ego-vehicle.
6. **non_ego_flow.X_relative_accelatation:** Calculated relative acceleration of all non-ego vehicles at each timestamp with the ego-vehicle.
7. **non_ego_flow.X _presence:** It represents the presence of that vehicle (0 or 1) at a particular timestamp. This additional column was introduced for NA values for vehicles that are not present in a particular timestamp. In this, we keep the relative acceleration, speed, and displacement as 0 and then add this additional column.
8. **Collision_occured:** 1 when a collision occurs and 0 when no collision occurs at a particular timestamp.

ARIMA: Univariate considering Ego-vehicle acceleration:

ARIMA is a forecasting technique that models a time series based on its lags and the lagged forecast errors. It's important to check if the time series is stationary i.e. mean, variance, and autocorrelation of the series, are all constant over time, as ARIMA works best with stationary data. For ARIMA we used one of the

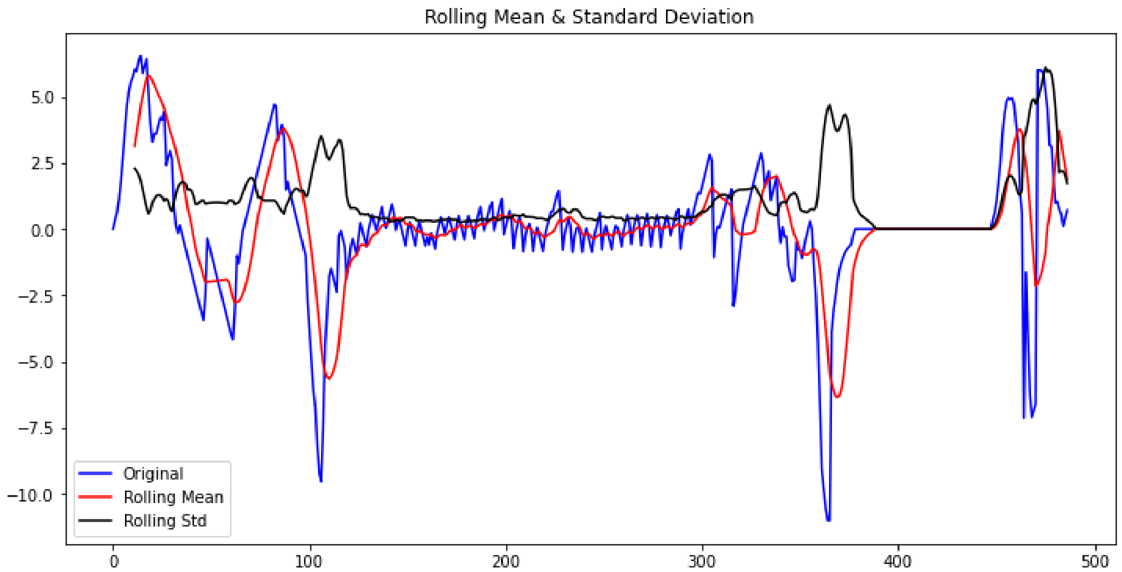
Checking the stationarity of the ego_acceleration time series:

```
Results of Dickey-Fuller Test:
Test Statistic                  -4.793910
p-value                          0.000056
#Lags Used                      10.000000
Number of Observations Used    476.000000
Critical Value (1%)             -3.444163
Critical Value (5%)             -2.867631
Critical Value (10%)            -2.570014
```

The results of the Dickey-Fuller test indicate that the ego_acceleration time series is stationary. This conclusion is based on the following observations:

The Test Statistic (-4.793910) is less than the Critical Value at 1% (-3.444163), 5% (-2.867631), and 10% (-2.570014) confidence levels.

The p-value (0.006745) is low, suggesting that the null hypothesis of the presence of a unit root can be rejected.



The overall mean and variance of the series do not remain constant over time, as seen by the red line (Rolling Mean) and black line(Rolling std), which varies with time. This suggests that the time series might be non-stationary. However, the statistical values suggest that the ego_acceleration time series is stationary. So we go with the statistical values that ego_acceleration is stationary.

Determining p-d-q value of ARIMA:

These parameters are p (order of the autoregressive part), d (degree of differencing), and q (order of the moving average part). I used Auto ARIMA, which automatically determines the best combination of p and q based on the AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion) values. d will be 0 the series is already stationary. From auto ARIMA we found the values ARIMA (3,0,2): p = 3, d = 0, and q = 2.

After applying an ARIMA model to forecast the value of ego_acceleration at the k+1 time step, the predicted acceleration for the 487th data point was 0.78 m/s^2. However, this prediction was inaccurate when compared to the actual recorded value of 0.41 m/s^2 at 49.2 seconds, which corresponds to the occurrence of a collision.